

RAW SEQUENCE LISTING

**The Biotechnology Systems Branch of the Scientific and Technical
Information Center (STIC) no errors detected.**

Application Serial Number: 10/686, 835A
Source: IFW16
Date Processed by STIC: 09/22/2005

ENTERED



IFW16

RAW SEQUENCE LISTING

DATE: 09/22/2005

PATENT APPLICATION: US/10/686,835A

TIME: 15:15:14

Input Set : A:\sequence listing.txt

Output Set: N:\CRF4\09222005\J686835A.raw

4 <110> APPLICANT: BOEHRINGER INGELHEIM (CANADA) LTD.
 6 <120> TITLE OF INVENTION: SELF REPLICATING RNA MOLECULE FROM
 7 HEPATITIS C VIRUS
 9 <130> FILE REFERENCE: 13/083
 C--> 11 <140> CURRENT APPLICATION NUMBER: US/10/686,835A
 C--> 11 <141> CURRENT FILING DATE: 2003-10-16
 11 <150> PRIOR APPLICATION NUMBER: 10/029,907
 12 <151> PRIOR FILING DATE: 2001-12-21
 14 <150> PRIOR APPLICATION NUMBER: 60/257,857
 15 <151> PRIOR FILING DATE: 2000-12-22
 17 <160> NUMBER OF SEQ ID NOS: 28
 19 <170> SOFTWARE: FastSEQ for Windows Version 4.0
 21 <210> SEQ ID NO: 1
 22 <211> LENGTH: 8639
 23 <212> TYPE: DNA
 24 <213> ORGANISM: HCV
 26 <220> FEATURE:
 27 <221> NAME/KEY: CDS
 28 <222> LOCATION: (1803)...(8408)
 30 <400> SEQUENCE: 1
 31 ggccagcccc cgattggggg cgacactcca ccatagatca ctccccctgtg aggaactact 60
 32 gtcttcacgc agaaagcgtc tagccatggc gttagtatga gtgtcgtgca gcctccagga 120
 33 cccccctcc cgaggagagcc atagtgggtct gcggaaccgg tgagtacacc ggaattgcca 180
 34 ggacgaccgg gtcctttctt ggatcaaccc gctcaatgcc tggagatttg ggcgtgcccc 240
 35 cgcgagactg ctagccgagt agtggtgggt cgcgaaaggc cttgtggtac tgcctgatag 300
 36 ggtgcttgcg agtgccccgg gaggtctcgt agaccgtgca ccatgagcac gaatcctaaa 360
 37 cctcaaagaa aaaccaaagg gcgcgccatg attgaacaag atggattgca cgcaggttct 420
 38 ccggccgctt ggggtggagag gctattcggc tatgactggg cacaacagac aatcggctgc 480
 39 tctgatgccg ccgtgttccg gctgtcagcg caggggcgcc cggttctttt tgtcaagacc 540
 40 gacctgtccg gtgccctgaa tgaactgcag gacgaggcag cgcggctatc gtggctggcc 600
 41 acgacggggc ttccttgccg agctgtgctc gacgttgctc ctgaagcggg aagggaactgg 660
 42 ctgctatttg gcgaagtgcc ggggcaggat ctctgtcat ctacacttgc tctgcccag 720
 43 aaagtatcca tcatggctga tgcaatgcgg cggctgcata cgcttgatcc ggctacctgc 780
 44 ccattcgacc accaagcgaa acatcgcac gagcgagcac gtactcggat ggaagccgg 840
 45 cttgtcgatc aggatgatct ggacgaagag catcaggggc tcgcgccagc cgaactgttc 900
 46 gccaggetca aggcgcgcac gcccgacggc gaggatctcg tcgtgaccca tggcgatgcc 960
 47 tgcttgccga atatcatggg ggaatggc cgcttttctg gattcatcga ctgtggccgg 1020
 48 ctgggtgtgg cggaccgcta tcaggacata gcgttggtc cccgtgatat tgctgaagag 1080
 49 cttggcggcg aatgggctga ccgcttcctc gtgctttacg gtatcgccgc tcccgaattc 1140
 50 cagcgcacgc ccttctatcg ccttcttgac gagttcttct gagttcgcgc ccagatgtta 1200
 51 acagaccaca acggtttccc tctagcggga tcaattccgc cccccccct aacgttactg 1260
 52 gccgaagccg cttggaataa ggccgggtgt cgtttgctc tatgttattt tccaccatat 1320
 53 tgccgtcttt tggcaatgtg agggcccgga aacctggccc tgtcttcttg acgagcattc 1380

RAW SEQUENCE LISTING

DATE: 09/22/2005

PATENT APPLICATION: US/10/686,835A

TIME: 15:15:14

Input Set : A:\sequence listing.txt

Output Set: N:\CRF4\09222005\J686835A.raw

```

54 ctagggggtct ttccctctc gccaaaggaa tgcaagggtct gttgaatgtc gtgaagggaag 1440
55 cagttcctctt ggaagcttct tgaagacaaa caacgtctgt agcgaccctt tgcaggcagc 1500
56 ggaaccccccc acctggcgac aggtgcctct gcggccaaaa gccacgtgta taagatacac 1560
57 ctgcaaaggc ggcacaaccc cagtgccacg ttgtgagttg gatagttgtg gaaagagtca 1620
58 aatggctctc ctcaagcgta ttcaacaagg ggctgaagga tgcccagaag gtacccatt 1680
59 gtatgggatac tgatctgggg cctcggtgca catgctttac atgtgtttag tgcagggttaa 1740
60 aaaacgtcta ggccccccga accacgggga cgtggttttc ctttgaaaaa cacgataata 1800
61 cc atg gac cgg gag atg gca gca tgc tgc gga ggc gcg gtt ttc gta 1847
62 Met Asp Arg Glu Met Ala Ala Ser Cys Gly Gly Ala Val Phe Val
63 1 5 10 15
65 ggt ctg ata ctc ttg acc ttg tca ccg cac tat aag ctg ttc ctc gct 1895
66 Gly Leu Ile Leu Leu Thr Leu Ser Pro His Tyr Lys Leu Phe Leu Ala
67 20 25 30
69 agg ctc ata tgg tgg tta caa tat ttt atc acc agg gcc gag gca cac 1943
70 Arg Leu Ile Trp Trp Leu Gln Tyr Phe Ile Thr Arg Ala Glu Ala His
71 35 40 45
73 ttg caa gtg tgg atc ccc ccc ctc aac gtt cgg ggg ggc cgc gat gcc 1991
74 Leu Gln Val Trp Ile Pro Pro Leu Asn Val Arg Gly Gly Arg Asp Ala
75 50 55 60
77 gtc atc ctc ctc acg tgc gcg atc cac cca gag cta atc ttt acc atc 2039
78 Val Ile Leu Leu Thr Cys Ala Ile His Pro Glu Leu Ile Phe Thr Ile
79 65 70 75
81 acc aaa atc ttg ctc gcc ata ctc ggt cca ctc atg gtg ctc cag gct 2087
82 Thr Lys Ile Leu Leu Ala Ile Leu Gly Pro Leu Met Val Leu Gln Ala
83 80 85 90 95
85 ggt ata acc aaa gtg ccg tac ttc gtg cgc gca cac ggg ctc att cgt 2135
86 Gly Ile Thr Lys Val Pro Tyr Phe Val Arg Ala His Gly Leu Ile Arg
87 100 105 110
89 gca tgc atg ctg gtg cgg aag gtt gct ggg ggt cat tat gtc caa atg 2183
90 Ala Cys Met Leu Val Arg Lys Val Ala Gly Gly His Tyr Val Gln Met
91 115 120 125
93 gct ctc atg aag ttg gcc gca ctg aca ggt acg tac gtt tat gac cat 2231
94 Ala Leu Met Lys Leu Ala Ala Leu Thr Gly Thr Tyr Val Tyr Asp His
95 130 135 140
97 ctc acc cca ctg cgg gac tgg gcc cac gcg ggc cta cga gac ctt gcg 2279
98 Leu Thr Pro Leu Arg Asp Trp Ala His Ala Gly Leu Arg Asp Leu Ala
99 145 150 155
101 gtg gca gtt gag ccc gtc gtc ttc tct gat atg gag acc aag gtt atc 2327
102 Val Ala Val Glu Pro Val Val Phe Ser Asp Met Glu Thr Lys Val Ile
103 160 165 170 175
105 acc tgg ggg gca gac acc gcg gcg tgt ggg gac atc atc ttg ggc ctg 2375
106 Thr Trp Gly Ala Asp Thr Ala Ala Cys Gly Asp Ile Ile Leu Gly Leu
107 180 185 190
109 ccc gtc tcc gcc cgc agg ggg agg gag ata cat ctg gga ccg gca gac 2423
110 Pro Val Ser Ala Arg Arg Gly Arg Glu Ile His Leu Gly Pro Ala Asp
111 195 200 205
113 agc ctt gaa ggg cag ggg tgg cga ctc ctc gcg cct att acg gcc tac 2471
114 Ser Leu Glu Gly Gln Gly Trp Arg Leu Leu Ala Pro Ile Thr Ala Tyr
115 210 215 220

```

RAW SEQUENCE LISTING

DATE: 09/22/2005

PATENT APPLICATION: US/10/686,835A

TIME: 15:15:14

Input Set : A:\sequence listing.txt

Output Set: N:\CRF4\09222005\J686835A.raw

```

117 tcc caa cag acg cga ggc cta ctt ggc tgc atc atc act agc ctc aca 2519
118 Ser Gln Gln Thr Arg Gly Leu Leu Gly Cys Ile Ile Thr Ser Leu Thr
119      225      230      235
121 ggc cgg gac agg aac cag gtc gag ggg gag gtc caa gtg gtc tcc acc 2567
122 Gly Arg Asp Arg Asn Gln Val Glu Gly Glu Val Gln Val Val Ser Thr
123 240      245      250      255
125 gca aca caa tct ttc ctg gcg acc tgc gtc aat ggc gtg tgt tgg act 2615
126 Ala Thr Gln Ser Phe Leu Ala Thr Cys Val Asn Gly Val Cys Trp Thr
127      260      265      270
129 gtc tat cat ggt gcc ggc tca aag acc ctt gcc ggc cca aag ggc cca 2663
130 Val Tyr His Gly Ala Gly Ser Lys Thr Leu Ala Gly Pro Lys Gly Pro
131      275      280      285
133 atc acc caa atg tac acc aat gtg gac cag gac ctc gtc ggc tgg caa 2711
134 Ile Thr Gln Met Tyr Thr Asn Val Asp Gln Asp Leu Val Gly Trp Gln
135      290      295      300
137 gcg ccc ccc ggg gcg cgt tcc ttg aca cca tgc acc tgc ggc agc tcg 2759
138 Ala Pro Pro Gly Ala Arg Ser Leu Thr Pro Cys Thr Cys Gly Ser Ser
139      305      310      315
141 gac ctt tac ttg gtc acg agg cat gcc gat gtc att ccg gtg cgc cgg 2807
142 Asp Leu Tyr Leu Val Thr Arg His Ala Asp Val Ile Pro Val Arg Arg
143 320      325      330      335
145 cgg ggc gac agc agg ggg agc cta ctc tcc ccc agg ccc gtc tcc tac 2855
146 Arg Gly Asp Ser Arg Gly Ser Leu Leu Ser Pro Arg Pro Val Ser Tyr
147      340      345      350
149 ttg aag ggc tct tcg ggc ggt cca ctg ctc tgc ccc tcg ggg cac gct 2903
150 Leu Lys Gly Ser Ser Gly Gly Pro Leu Leu Cys Pro Ser Gly His Ala
151      355      360      365
153 gtg ggc atc ttt cgg gct gcc gtg tgc acc cga ggg gtt gcg aag gcg 2951
154 Val Gly Ile Phe Arg Ala Ala Val Cys Thr Arg Gly Val Ala Lys Ala
155      370      375      380
157 gtg gac ttt gta ccc gtc gag tct atg gaa acc act atg cgg tcc ccg 2999
158 Val Asp Phe Val Pro Val Glu Ser Met Glu Thr Thr Met Arg Ser Pro
159      385      390      395
161 gtc ttc acg gac aac tcg tcc cct ccg gcc gta ccg cag aca ttc cag 3047
162 Val Phe Thr Asp Asn Ser Ser Pro Pro Ala Val Pro Gln Thr Phe Gln
163 400      405      410      415
165 gtg gcc cat cta cac gcc cct act ggt agc ggc aag agc act aag gtg 3095
166 Val Ala His Leu His Ala Pro Thr Gly Ser Gly Lys Ser Thr Lys Val
167      420      425      430
169 ccg gct gcg tat gca gcc caa ggg tat aag gtg ctt gtc ctg aac ccg 3143
170 Pro Ala Ala Tyr Ala Ala Gln Gly Tyr Lys Val Leu Val Leu Asn Pro
171      435      440      445
173 tcc gtc gcc gcc acc cta ggt ttc ggg gcg tat atg tct aag gca cat 3191
174 Ser Val Ala Ala Thr Leu Gly Phe Gly Ala Tyr Met Ser Lys Ala His
175      450      455      460
177 ggt atc gac cct aac atc aga acc ggg gta agg acc atc acc acg ggt 3239
178 Gly Ile Asp Pro Asn Ile Arg Thr Gly Val Arg Thr Ile Thr Thr Gly
179      465      470      475
181 gcc ccc atc acg tac tcc acc tat ggc aag ttt ctt gcc gac ggt ggt 3287

```

RAW SEQUENCE LISTING

DATE: 09/22/2005

PATENT APPLICATION: US/10/686,835A

TIME: 15:15:14

Input Set : A:\sequence listing.txt

Output Set: N:\CRF4\09222005\J686835A.raw

182	Ala	Pro	Ile	Thr	Tyr	Ser	Thr	Tyr	Gly	Lys	Phe	Leu	Ala	Asp	Gly	Gly	
183	480					485					490					495	
185	tgc	tct	ggg	ggc	gcc	tat	gac	atc	ata	ata	tgt	gat	gag	tgc	cac	tca	3335
186	Cys	Ser	Gly	Gly	Ala	Tyr	Asp	Ile	Ile	Ile	Cys	Asp	Glu	Cys	His	Ser	
187					500					505					510		
189	act	gac	tcg	acc	act	atc	ctg	ggc	atc	ggc	aca	gtc	ctg	gac	caa	gcg	3383
190	Thr	Asp	Ser	Thr	Thr	Ile	Leu	Gly	Ile	Gly	Thr	Val	Leu	Asp	Gln	Ala	
191					515					520					525		
193	gag	acg	gct	gga	gcg	cga	ctc	gtc	gtg	ctc	gcc	acc	gct	acg	cct	ccg	3431
194	Glu	Thr	Ala	Gly	Ala	Arg	Leu	Val	Val	Leu	Ala	Thr	Ala	Thr	Pro	Pro	
195					530					535					540		
197	gga	tcg	gtc	acc	gtg	cca	cat	cca	aac	atc	gag	gag	gtg	gct	ctg	tcc	3479
198	Gly	Ser	Val	Thr	Val	Pro	His	Pro	Asn	Ile	Glu	Glu	Val	Ala	Leu	Ser	
199					545					550					555		
201	agc	act	gga	gaa	atc	ccc	ttt	tat	ggc	aaa	gcc	atc	ccc	atc	gag	acc	3527
202	Ser	Thr	Gly	Glu	Ile	Pro	Phe	Tyr	Gly	Lys	Ala	Ile	Pro	Ile	Glu	Thr	
203	560					565					570					575	
205	atc	aag	ggg	ggg	agg	cac	ctc	att	ttc	tgc	cat	tcc	aag	aag	aaa	tgt	3575
206	Ile	Lys	Gly	Gly	Arg	His	Leu	Ile	Phe	Cys	His	Ser	Lys	Lys	Lys	Cys	
207					580						585					590	
209	gat	gag	ctc	gcc	gcg	aag	ctg	tcc	ggc	ctc	gga	ctc	aat	gct	gta	gca	3623
210	Asp	Glu	Leu	Ala	Ala	Lys	Leu	Ser	Gly	Leu	Gly	Leu	Asn	Ala	Val	Ala	
211					595					600					605		
213	tat	tac	cgg	ggc	ctt	gat	gta	tcc	gtc	ata	cca	act	agc	gga	gac	gtc	3671
214	Tyr	Tyr	Arg	Gly	Leu	Asp	Val	Ser	Val	Ile	Pro	Thr	Ser	Gly	Asp	Val	
215					610					615					620		
217	att	gtc	gta	gca	acg	gac	gct	cta	atg	acg	ggc	ttt	acc	ggc	gat	ttc	3719
218	Ile	Val	Val	Ala	Thr	Asp	Ala	Leu	Met	Thr	Gly	Phe	Thr	Gly	Asp	Phe	
219					625						630				635		
221	gac	tca	gtg	atc	gac	tgc	aat	aca	tgt	gtc	acc	cag	aca	gtc	gac	ttc	3767
222	Asp	Ser	Val	Ile	Asp	Cys	Asn	Thr	Cys	Val	Thr	Gln	Thr	Val	Asp	Phe	
223	640					645					650					655	
225	agc	ctg	gac	ccg	acc	ttc	acc	att	gag	acg	acc	gtg	cca	caa	gac		3815
226	Ser	Leu	Asp	Pro	Thr	Phe	Thr	Ile	Glu	Thr	Thr	Val	Pro	Gln	Asp		
227					660						665				670		
229	gcg	gtg	tca	cgc	tcg	cag	cgg	cga	ggc	agg	act	ggt	agg	ggc	agg	atg	3863
230	Ala	Val	Ser	Arg	Ser	Gln	Arg	Arg	Gly	Arg	Thr	Gly	Arg	Gly	Arg	Met	
231					675					680					685		
233	ggc	att	tac	agg	ttt	gtg	act	cca	gga	gaa	cgg	ccc	tcg	ggc	atg	ttc	3911
234	Gly	Ile	Tyr	Arg	Phe	Val	Thr	Pro	Gly	Glu	Arg	Pro	Ser	Gly	Met	Phe	
235					690					695					700		
237	gat	tcc	tcg	gtt	ctg	tgc	gag	tgc	tat	gac	gcg	ggc	tgt	gct	tgg	tac	3959
238	Asp	Ser	Ser	Val	Leu	Cys	Glu	Cys	Tyr	Asp	Ala	Gly	Cys	Ala	Trp	Tyr	
239					705					710					715		
241	gag	ctc	acg	ccc	gcc	gag	acc	tca	gtt	agg	ttg	cgg	gct	tac	cta	aac	4007
242	Glu	Leu	Thr	Pro	Ala	Glu	Thr	Ser	Val	Arg	Leu	Arg	Ala	Tyr	Leu	Asn	
243	720					725					730					735	
245	aca	cca	ggg	ttg	ccc	gtc	tgc	cag	gac	cat	ctg	gag	ttc	tgg	gag	agc	4055
246	Thr	Pro	Gly	Leu	Pro	Val	Cys	Gln	Asp	His	Leu	Glu	Phe	Trp	Glu	Ser	

RAW SEQUENCE LISTING

DATE: 09/22/2005

PATENT APPLICATION: US/10/686,835A

TIME: 15:15:14

Input Set : A:\sequence listing.txt

Output Set: N:\CRF4\09222005\J686835A.raw

247		740		745		750	
249	gtc ttt aca ggc ctc acc cac ata gac gcc cat ttc ttg tcc cag act						4103
250	Val Phe Thr Gly Leu Thr His Ile Asp Ala His Phe Leu Ser Gln Thr						
251		755		760		765	
253	aag cag gca gga gac aac ttc ccc tac ctg gta gca tac cag gct acg						4151
254	Lys Gln Ala Gly Asp Asn Phe Pro Tyr Leu Val Ala Tyr Gln Ala Thr						
255		770		775		780	
257	gtg tgc gcc agg gct cag gct cca cct cca tcg tgg gac caa atg tgg						4199
258	Val Cys Ala Arg Ala Gln Ala Pro Pro Pro Ser Trp Asp Gln Met Trp						
259		785		790		795	
261	aag tgt ctc ata cgg cta aag cct acg ctg cac ggg cca acg ccc ctg						4247
262	Lys Cys Leu Ile Arg Leu Lys Pro Thr Leu His Gly Pro Thr Pro Leu						
263	800		805		810		815
265	ctg tat agg ctg gga gcc gtt caa aac gag gtt act acc aca cac ccc						4295
266	Leu Tyr Arg Leu Gly Ala Val Gln Asn Glu Val Thr Thr Thr His Pro						
267		820		825		830	
269	ata acc aaa tac atc atg gca tgc atg tcg gct gac ctg gag gtc gtc						4343
270	Ile Thr Lys Tyr Ile Met Ala Cys Met Ser Ala Asp Leu Glu Val Val						
271		835		840		845	
273	acg agc acc tgg gtg ctg gta ggc gga gtc cta gca gct ctg gcc gcg						4391
274	Thr Ser Thr Trp Val Leu Val Gly Gly Val Leu Ala Ala Leu Ala Ala						
275		850		855		860	
277	tat tgc ctg aca aca ggc agc gtg gtc att gtg ggc agg atc atc ttg						4439
278	Tyr Cys Leu Thr Thr Gly Ser Val Val Ile Val Gly Arg Ile Ile Leu						
279		865		870		875	
281	tcc gga aag ccg gcc atc att ccc gac agg gaa gtc ctt tac cgg gag						4487
282	Ser Gly Lys Pro Ala Ile Ile Pro Asp Arg Glu Val Leu Tyr Arg Glu						
283	880		885		890		895
285	ttc gat gag atg gaa gag tgc gcc tca cac ctc cct tac atc gaa cag						4535
286	Phe Asp Glu Met Glu Glu Cys Ala Ser His Leu Pro Tyr Ile Glu Gln						
287		900		905		910	
289	gga atg cag ctc gcc gaa caa ttc aaa cag aag gca atc ggg ttg ctg						4583
290	Gly Met Gln Leu Ala Glu Gln Phe Lys Gln Lys Ala Ile Gly Leu Leu						
291		915		920		925	
293	caa aca gcc acc aag caa gcg gag gct gct gct ccc gtg gtg gaa tcc						4631
294	Gln Thr Ala Thr Lys Gln Ala Glu Ala Ala Ala Pro Val Val Glu Ser						
295		930		935		940	
297	aag tgg cgg acc ctc gaa gcc ttc tgg gcg aag cat atg tgg aat ttc						4679
298	Lys Trp Arg Thr Leu Glu Ala Phe Trp Ala Lys His Met Trp Asn Phe						
299		945		950		955	
301	atc agc ggg ata caa tat tta gca ggc ttg tcc act ctg cct ggc aac						4727
302	Ile Ser Gly Ile Gln Tyr Leu Ala Gly Leu Ser Thr Leu Pro Gly Asn						
303	960		965		970		975
305	ccc gcg ata gca tca ctg atg gca ttc aca gcc tct atc acc agc ccg						4775
306	Pro Ala Ile Ala Ser Leu Met Ala Phe Thr Ala Ser Ile Thr Ser Pro						
307		980		985		990	
309	ctc acc acc caa cat acc ctc ctg ttt aac atc ctg ggg gga tgg gtg						4823
310	Leu Thr Thr Gln His Thr Leu Leu Phe Asn Ile Leu Gly Gly Trp Val						
311		995		1000		1005	

RAW SEQUENCE LISTING ERROR SUMMARY
PATENT APPLICATION: US/10/686,835A

DATE: 09/22/2005
TIME: 15:15:15

Input Set : A:\sequence listing.txt
Output Set: N:\CRF4\09222005\J686835A.raw

Please Note:

Use of n and/or Xaa have been detected in the Sequence Listing. Please review the Sequence Listing to ensure that a corresponding explanation is presented in the <220> to <223> fields of each sequence which presents at least one n or Xaa.

Seq#:2; Xaa Pos. 882,1489

Seq#:3; Xaa Pos. 882,1489

VERIFICATION SUMMARY

DATE: 09/22/2005

PATENT APPLICATION: US/10/686,835A

TIME: 15:15:15

Input Set : A:\sequence listing.txt

Output Set: N:\CRF4\09222005\J686835A.raw

L:11 M:270 C: Current Application Number differs, Replaced Current Application No
L:11 M:271 C: Current Filing Date differs, Replaced Current Filing Date
L:631 M:258 W: Mandatory Feature missing, <220> Tag not found for SEQ ID#:2
L:635 M:258 W: Mandatory Feature missing, <220> Tag not found for SEQ ID#:2
L:887 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:2 after pos.:4489
M:341 Repeated in SeqNo=2
L:1237 M:258 W: Mandatory Feature missing, <220> Tag not found for SEQ ID#:3
L:1348 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:3 after pos.:880
M:341 Repeated in SeqNo=3